

Developing Spatial Surrogates for Modelling Applications

J. Wayne Boulton*
Rowan Williams Davies and Irwin Inc. (RWDI)
650 Woodlawn Rd. West
Guelph, Ontario, Canada, N1K 1B8
wb@rwdi.com

Kristofer A. Siriunas
RWDI West Inc.
#1800, 840-7th Avenue S.W.
Calgary, Alberta, Canada, T2P 3G2

Mike Lepage, Stephanie Schill
Rowan Williams Davies and Irwin Inc. (RWDI)
650 Woodlawn Rd. West
Guelph, Ontario, Canada, N1K 1B8

Session: *Emission Inventory Preparation for Modeling*

ABSTRACT

There are a number of programs and modelling systems available for processing regional emission inventory data for modelling applications. These models are used to temporalize, speciate and spatially allocate State- or County-wide emissions to a model grid for photochemical or dispersion modelling. The spatial allocation of emissions is done using spatial surrogates which normally take the form of digital maps of road networks or census data. The preparation of spatial surrogates is typically cumbersome due to specific software needs and the inherent complexity in modifying the emission processors. A need to produce spatial surrogates faster using readily available information was identified. This paper provides a review of techniques for preparing spatial surrogates using cheaper, more readily available software and data. ArcView and ARC/INFO commercial GIS (Geographic Information System) software packages were used to develop unique programs and tools for this purpose. These programs were written to allow for the inclusion of any type or number of spatial surrogates; from traditional sources such as road networks and census data, to less traditional sources such as airports and flight paths, marine transport routes, railways, industrial, commercial, and residential land uses, and so on. These techniques and software tools have been used with great success in the preparation of spatial surrogate ratios for use with emission models such as SMOKE, as well as in the apportionment of emissions data into gridded area sources for use with dispersion models such as CALPUFF and AERMOD.

INTRODUCTION

The primary goal of most air quality modelling applications is to accurately replicate the physical and chemical processes that occur over a specific study area or region in order to arrive at an accurate simulation of the transport and dispersion of airborne pollutants. Inherent in this is the desire to replicate the spatial distribution of emissions throughout the study area. In an ideal situation, the physical location and release characteristics of all emissions would be known exactly. However, the spatial allocation of emissions in a modelling inventory only approximate the actual location of the emission sources at best (US EPA, 2002).

Spatial surrogates are used to allocate geographically distributed data to higher resolution (i.e., smaller) geographic areas based on some form of activity or socio-economic / demographic data. In the context of air quality modelling, spatial surrogates are used to allocate emissions that are usually provided on a State- or County-wide basis, to smaller geo-political units such as enumeration areas or census tracts, or to a model grid for specific modelling purposes.

The three fundamental components to the spatial surrogate process are:

1. The raw data to be spatially allocated (i.e., emission inventory data, typically produced at either the State- or County-wide level);
2. The spatial surrogate itself, which typically takes the form of a digital map representing some form of activity or socio-economic / demographic data (i.e., census information, road network, etc.); and,
3. A higher resolution geo-political map or model grid to which the emission data will be allocated using the spatial surrogate.

Although not discussed in detail in this paper, an integral part of the spatial allocation process is the development of linkages between the surrogates and the emission sources themselves. To accomplish this step, cross-reference files are created in which each emission source (typically defined by unique Source Classification Code or SCC) is assigned a specific spatial surrogate. For example, a roadway surrogate developed using a digital road network would be used to apportion all emissions attributed to on-road emission sources. In addition, depending on the emission model being used, different surrogates can be used for the same source type in different States or Counties, depending on the availability of information.

Improvements in modelling capabilities through the advent of cheaper, more readily available computing power over the past few years has allowed for the ability to perform more complex, higher resolution air quality dispersion and photochemical modelling. Inherent in this increase in modelling capability is the desire to improve the accuracy of how emissions are spatially represented within the various models. The US EPA provides guidance to air quality modellers on the selection of spatial allocation procedures in a number of recent documents (Stella, 2002; US EPA, 2002). In addition, several recent studies have also been performed with the common goal of finding better techniques for allocating emissions to a model grid, particularly for the allocation of on-road vehicular emissions (e.g., Kinnee *et al.*, 2001).

Historically, the preparation of spatial surrogates has been rather cumbersome due to specific hardware and software needs and the inherent complexity in modifying the emission processors themselves. A need to produce spatial surrogates for modelling applications using readily available data and less expensive computer software and hardware systems was identified.

This paper presents examples from three case studies pertaining to different techniques used to develop spatial surrogates for modelling applications. These case studies also outline new ways in which the surrogate data are being used for a variety of non-traditional modelling applications.

Case Study 1: PATH Modelling of the Pearl River Delta region, China and Hong Kong

An extensive modelling study was performed using the Hong Kong Environmental Protection Department's PATH model (Pollutants in the Atmosphere and their Transport over Hong Kong). The underlying purpose of the study was to perform regional air quality and smog modelling over the Hong Kong Special Administrative Region (HK SAR) and the Pearl River Delta Economic Zone (PRDEZ) and expand and improve upon the original PATH model resolution over the major urban centres of the PRDEZ, namely: Hong Kong, Shenzhen, Guangzhou, Zhuhai, Foshan, Dongguan, and Macau. The pollutants of interest included: respirable particulate matter (RSP), ozone, and nitrogen dioxide (NO₂).

For this modelling application, nested model domains were configured within PATH with the following horizontal resolutions: 40.5, 13.5, 4.5, and 1.5. The largest model domain covers Southern Mainland China, extending as far as Calcutta to the west, the Philippines to the east, Shanghai to the north, and beyond Vietnam to the south. In the present study, the original 1.5 km grid (covering only Hong Kong) was replaced with a new, expanded model domain covering a 21,000 km² area of the Pearl River Delta at a resolution of 1.5 km by 1.5 km.

Emission processing was performed using a modified version of EMS-95 on a Silicon Graphics (SGI) UNIX server and required the acquisition of UNIX versions of both SAS and ARC/INFO, both of which are costly third party software packages. EMS-95 was used to produce hourly gridded and speciated emission data for eight of eleven historical meteorological episodes modelled in MM5 (Mesoscale Meteorological Model 5). The output from EMS-95 was used for photochemical modelling with a modified version of SAQM (SARMAP Air Quality Model) developed specifically for PATH.

As part of this study, a base year emission inventory for 1997 was developed for the PRDEZ and the existing 1995 inventory for Hong Kong was projected to 1997. The 1997 inventories were also projected to four future years: 2000, 2005, 2010 and 2015, although only the Base Year (1997) and 2010 future year were modelled in the air quality simulations. A future year (2010) business-as-usual (BAU) growth scenario and two future year emission control scenarios were modelled.

Annual emissions were calculated by activity type (i.e., SCC or ASCT code), and by "State" or similar geo-political unit. These State-wide emissions were then allocated to individual "Counties" based on available demographic and activity information. Census data for 1996 were used to allocate Hong Kong wide emissions to 18 "Counties". Data from the 1998 Statistical Yearbook of Guangdong were used to allocate State-wide emissions to their respective "Counties" throughout the Chinese Province of Guangdong.

To allocate emissions to the new 1.5 km grid, new spatial surrogates were required. Although the GIS data used to create spatial surrogates for Hong Kong were relatively up to date, GIS data for the Chinese portion of the model domain were limited, requiring the development of new GIS coverages and the manual manipulation of existing data sources to arrive at new surrogates for this part of the domain.

The following is a partial list of key emission sources for which surrogates were developed in this study:

- Marine Vessels

Emissions from marine sources were allocated to the model grid using polygon area definitions for each of the major harbour areas: Macau Harbour, Hong Kong Harbour and moorings, Hong Kong SAR waters, Pearl River Delta waters and Pearl River Estuary, and international waters. GIS polyline files depicting detailed ferry and marine transport routes within Hong Kong waters were also used to apportion marine emissions within the HK SAR waters.

- Area Sources

A digital land use coverage based on high resolution satellite imagery was provided and incorporated into the MM5 modelling. Figure 1 depicts the land use data gridded for the expanded 1.5 km MM5 model domain. This same coverage was also modified to create urban and rural land use surrogates for allocating area source emissions to the model grid. Areas classified as cultivation were used to apportion emissions from agricultural activities (e.g., agricultural waste burning, grain drying, pesticide application, etc.) in both the PRDEZ and Hong Kong. Areas defined as urban were used to apportion emissions from human activities (e.g., solvent use, construction dust, residential fuel combustion, etc.) in the HK SAR. For the PRDEZ, a census coverage was developed in ARC/INFO using urban and rural population statistics from

the 1998 Statistical Yearbook of Guangdong in combination with a GIS coverage of County boundaries for that region. This was then used to create a population based surrogate which in turn was used to apportion emissions from human activities to the model grid.

- **On-Road Vehicular Sources**

On-road vehicle emissions in Hong Kong were allocated to 18 counties based on available vehicle activity data. A digital road network for Hong Kong provided in the original PATH model was then used to apportion these County-wide emissions to the model grid. For mainland Guangdong Province, a coarse digital road network was provided in the original PATH model. However, this road network provided the location of the major highways only and contained no information on urban road networks (e.g., roads in Guangzhou, Foshan, Zhongshan, etc.) where vehicular emissions are typically very high. Using the GIS coverage to create an “as-is” mobile surrogate would have resulted in all vehicle emissions being apportioned to major highways outside of the urban centres, with no emissions from mobile sources emanating from within the urban areas. To address this issue the highway road network was converted into a polygon coverage using a buffer operation and then combined (i.e., unioned) with those areas defined as urban from the high resolution land use data. The end product of this procedure was a hybrid (i.e., combined urban land use and digital road network) road surrogate and is presented in Figure 2.

- **Airports**

Five major airports were identified in the emission inventory for which spatial surrogates were required. In Hong Kong, the base year 1997 airport emissions were assigned to the old Kai Tak Airport; whereas for all future year scenarios, airport emissions were apportioned to the new Chek Lap Kok Airport due to the transfer of airport activity to the newly constructed airport in 1998. Figure 3 depicts the spatial surrogates used to apportion emissions to the model grid for the Kai Tak, Chek Lap Kok and Macau International airports. Unique geographic areas defining emission activities were created for aircraft climbout (red), idling / taxiing (pink), takeoff (light blue), and approach / landing (dark blue). A general airport footprint was used to allocate emissions from ground support equipment and related activities.

Detailed results from the air quality modelling performed with SAQM are currently being reviewed by the contracting authority. Preliminary results indicate that the expanded 1.5 km domain and improvement to existing surrogates, as well as the creation of new surrogates for parts of China, allowed for the better apportionment of emissions to the model grid and hence a more representative distribution of emissions and resultant pollutant concentrations in both the PRDEZ and Hong Kong.

Having said this, the overall spatial surrogate and related modelling approach adopted in this study was very cumbersome. Data acquisition was difficult and in some cases, up to date GIS data were simply not available. The creation of the surrogates required the manipulation of EMS-95 scripts on a UNIX platform, creating the need to have in-house expertise in UNIX systems, ARCINFO, and SAS, as well as their respective macro languages (e.g., AML).

Province of Alberta, Alberta, Canada

In 2000, RWDI was retained by Environment Canada to prepare spatial surrogates for mobile and area emission sources for a model grid covering the entire province of Alberta, Canada. The SMOKE (Sparse Matrix Operating Kernel Emissions) emission processor produced by MCNC was selected by Environment Canada as the emission model of choice to prepare the required inputs for planned UAM photochemical model applications. In this study, the spatial surrogates used to allocate area source emissions to the model grid were essentially the same as those originally used by the Pollution Data Branch of Environment Canada to allocate province-wide emissions to census divisions, with a few minor modifications. The emission inventory in question is Environment Canada's 1995 Criteria Air Contaminant

(CAC) Emission Inventory.

The grid spacing of interest for this study was defined as 0.05° longitude by 0.025° latitude over a domain covering the entire province of Alberta (approximately 1,200 km north-south by 600 km east-west) and consisting of 88,000 grid cells.

The spatial surrogate file format required for input into SMOKE is relatively simple. Because of the flexibility within SMOKE to handle any number / type of spatial surrogates, a new approach was adopted to develop surrogates for this model application. ESRI's ARCVIEW GIS software (ver. 3.2a) was selected over the traditional ARC/INFO software in which to perform the overlay and surrogate calculation procedures. Scripts were written in Avenue (the resident, object oriented scripting language in ARCVIEW) which allowed for the calculation of surrogate ratios by grid cell over the entire domain.

Census division emissions are allocated to the grid cells in SMOKE using surrogate ratios. A surrogate ratio is the ratio of the total surrogate data within a grid cell divided by the total surrogate within the census division in which the grid cell is located. For example, if a digital road network is being used to create a spatial surrogate for "primary roads", then the surrogate ratio for an arbitrary grid cell is equal to the total length of all primary roads in the grid cell divided by the total length of all primary roads in the entire census division.

As part of this new approach, significant flexibility was added to the surrogate creation process. In addition to the ability to create more common spatial surrogates (i.e., population, road network, rail network, etc.), the ability to use different socio-economic and demographic parameters based on enumeration areas was also incorporated. In this way it was possible to create many surrogates in fewer steps using the different statistics from the enumeration area based census data (e.g., population, housing, total agricultural and related service industries, total logging and forest industries, total manufacturing industries, etc.). Figures 4 and 5 show the resultant gridded roadway and population spatial surrogates, respectively.

In addition to the development of the surrogates, Environment Canada was also interested in producing maps of gridded annual emissions based on the newly created surrogates. This too was readily performed in ARCVIEW. Because the surrogate data were provided by census division, it was possible to establish the required linkages between the surrogate and 1995 CAC emission inventory databases through ARCVIEW, which in turn allowed for the preparation of the required maps of gridded annual emissions. Figures 6 and 7 depict the gridded annual emissions from area and mobile emissions (combined) for NO_x and SO_x, respectively.

Over 20 different spatial surrogates were created using this newly devised approach. This normally complicated and data intensive process was made relatively simple using ARCVIEW. In addition, because ARCVIEW is a relatively inexpensive and readily available software package on PC, the routines developed in this project are readily adaptable to other applications. Having completed the Alberta Emission Inventory project, RWDI was later approached by Environment Canada to re-create the same spatial surrogates for a different grid structure, based on a different projection system and 4 km by 4 km grid spacing. These new data are to be used with the CMAQ (Community Scale Air Quality Model used in Models-3) as opposed to UAM. The original scripts were easily adapted for the new domain and grid spacing, and the surrogates were re-created with minimal effort.

City of Toronto, Ontario, Canada

Since the development of this new methodology for processing and manipulating spatial surrogates, two other air quality modelling studies have been able to take advantage of this unique approach.

In 2000, RWDI was retained by the City of Toronto, Ontario, Canada to develop a GIS software tool that would allow the City staff to model emissions and the resultant transport of primary pollutants across the City and to identify local hot spots. The modelling was based on Environment Canada's 1995 CAC Emission Inventory. To accomplish this task, a set of scripts were developed to create gridded spatial surrogates based on a detailed road network (for on-road mobile sources) and land use zoning information (for area emission sources). In addition, the top 50 point sources (based on combined $\text{SO}_x + \text{NO}_x$ emissions) were isolated and modelled as discrete points. The remainder of the point sources were allocated to the model grid and aggregated into the area source emissions. Due to runtime limitations within CALPUFF and the total number of sources being modelled, the model grid spacing was defined as 2 km by 2 km. Dispersion modelling was performed using CALPUFF, with gridded emissions treated as area sources within CALPUFF and point sources being modelled as discrete points.

The software tool was developed entirely in ARCVIEW. The tool allows for the modification of land use classes, changes in road network, etc. to be made in ARCVIEW, after which the menu-driven scripts can be used to re-run the spatial surrogate processors. The end result of this process is a CALPUFF input control file containing the newly updated source information which can then be run and used to map resultant pollutant concentrations across the City.

Greater Toronto Airports Authority, Toronto, Ontario, Canada

A similar project to that performed for the City of Toronto is currently in progress for the Greater Toronto Airports Authority (GTAA). As part of an ongoing airport expansion project, a detailed air quality assessment is being performed for Lester B. Pearson International Airport using emissions developed within EDMS (US FAA Emissions and Dispersion Modeling System) and dispersion modelling being performed using AERMOD.

As part of this assessment, it was deemed necessary to quantify the impact from background emissions attributed to sources within a 7.5 km radius of the airport. To develop the required gridded emissions for sources located within the study area, the same spatial allocation techniques and ARCVIEW scripts used in the City of Toronto study were modified and adopted in this study for the unique domain and raw GIS data.

Gridded digital land use and road network data were used to allocate emissions to a coarse model grid for input into AERMOD within ARCVIEW. Similarly, the top 15 point sources (based on combined $\text{SO}_x + \text{NO}_x$ emissions) were entered into AERMOD as discrete point sources, with the remainder of the point sources being aggregated into the gridded area sources. Figure 8 shows the overall location of LBPIA International Airport, the emission inventory study area, and the digital road network used to apportion on-road emissions to the model grid for dispersion modelling in AERMOD.

CONCLUSIONS

As noted earlier, the primary goal of most air quality modelling applications is to accurately replicate the physical and chemical processes that occur over a specific study area or region in order to arrive at an accurate simulation of the transport and dispersion of airborne pollutants. Inherent in this is the desire to replicate the spatial distribution of emissions throughout the study area. Through the development of new spatial allocation techniques and the creation of scripts using cheaper, more readily available GIS software on a PC platform, we are able to produce unique spatial surrogates for a variety of model applications and domains in a shorter amount of time than previously capable.

REFERENCES

United States Environmental Protection Department (U.S. EPA), 2002, *Spatial Allocation*, <www.epa.gov/ttnchie1/emch/spatial/index.html>, revision March 12, 2002.

Stella G. (U.S. EPA), 2002, Emission Factor and Inventory Group, Re: Spatial Allocation Files Located on EMCH, revision March 5, 2002, <http://www.epa.gov/ttn/chief/emch/spatial/emch_spatial_data.pdf>.

Ng, Y.L. and R.C. Joynt, 2001, "A PC-Based Emission Database System for the Auckland Region", In *Proceedings: Emission Inventory Conference: One Atmosphere, One Inventory, Many Challenges*, May 1 - 3, 2001.

Kinnee, E. Beidler, A., Possiel, N., and G. Stella, 2001, "Revised Methodology for the Spatial Allocation of VMT and Mobile Source Emissions Data", In *Proceedings: Emission Inventory Conference: One Atmosphere, One Inventory, Many Challenges*, May 1 - 3, 2001.

KEY WORDS

Spatial Surrogates
Spatial Allocation
Emission Inventory
SMOKE
ArcView
GIS

Figure 1. Gridded land use data used in MM5 and in the creation of spatial surrogates.

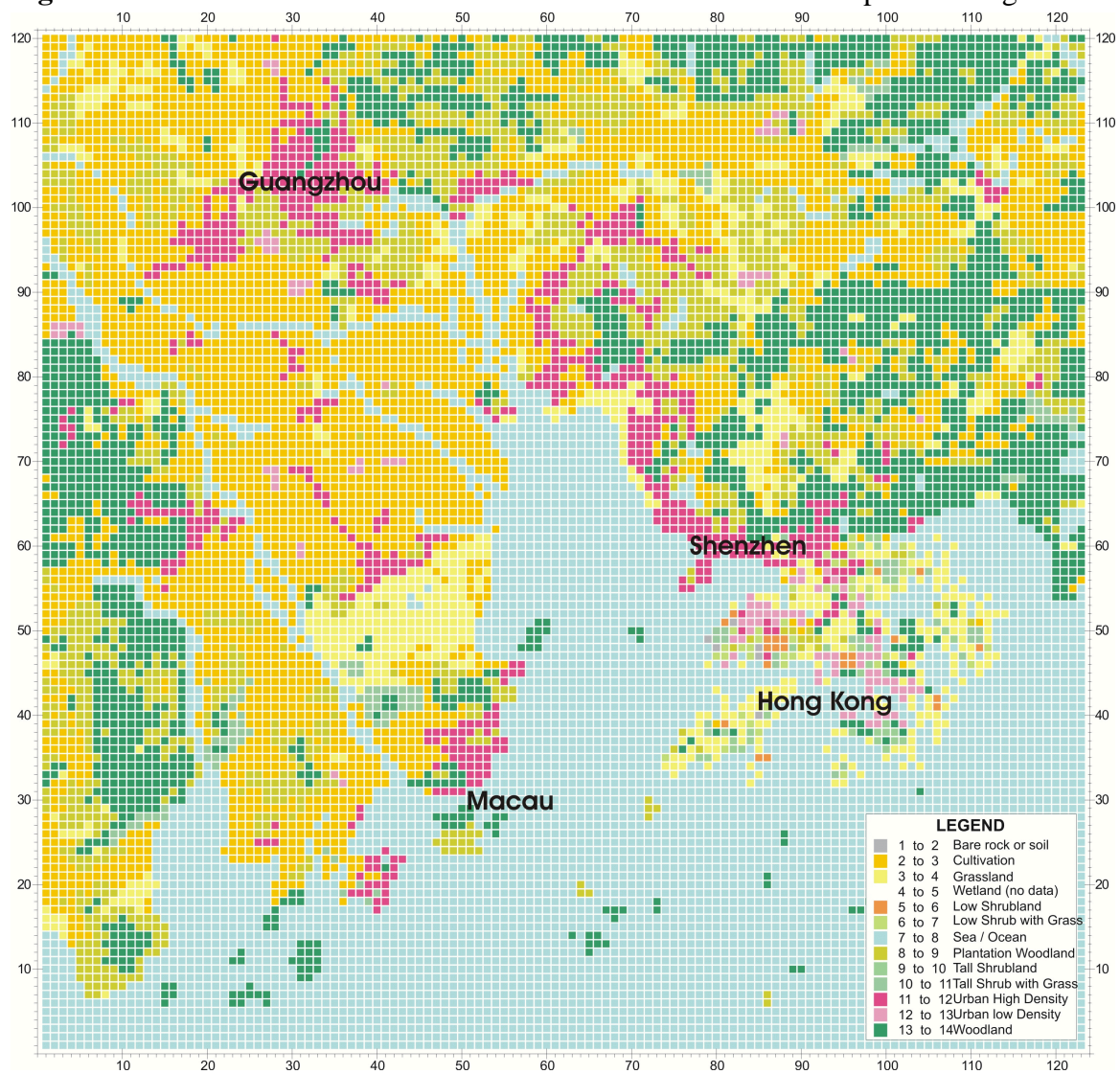


Figure 2. Combined PRDEZ urban + road surrogate.

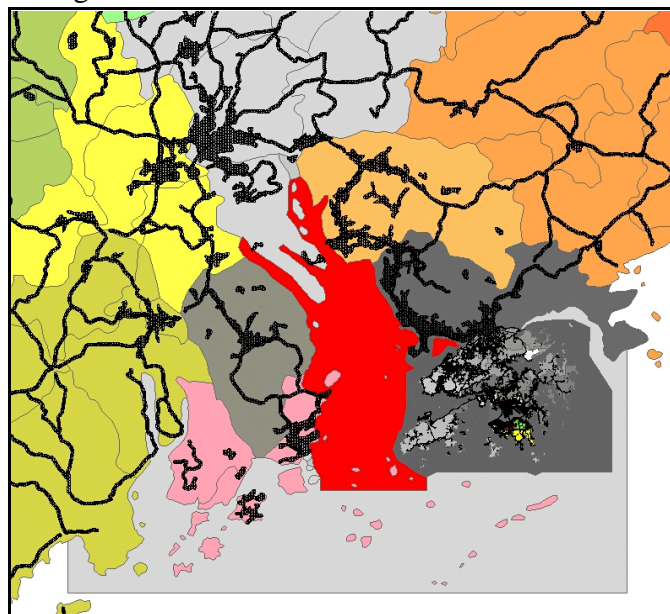


Figure 3. Airport spatial surrogates.

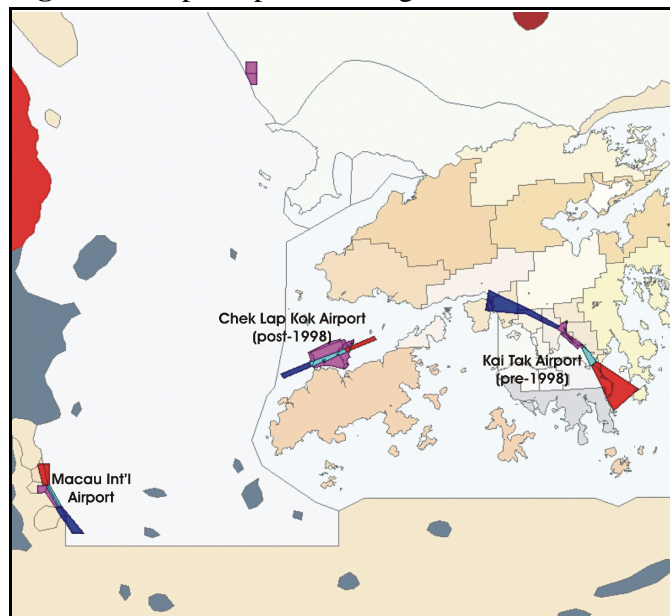


Figure 4. Gridded Alberta population surrogate ratios.

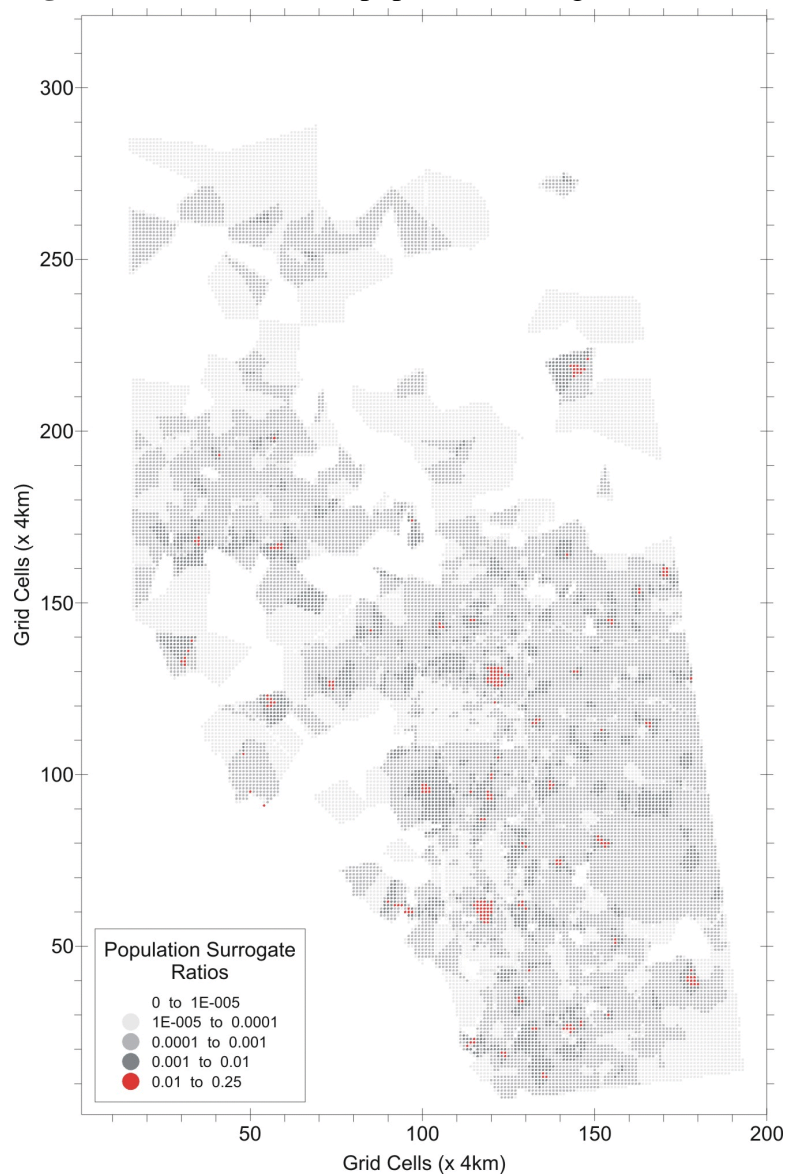


Figure 5. Gridded Alberta roadway surrogate ratios.

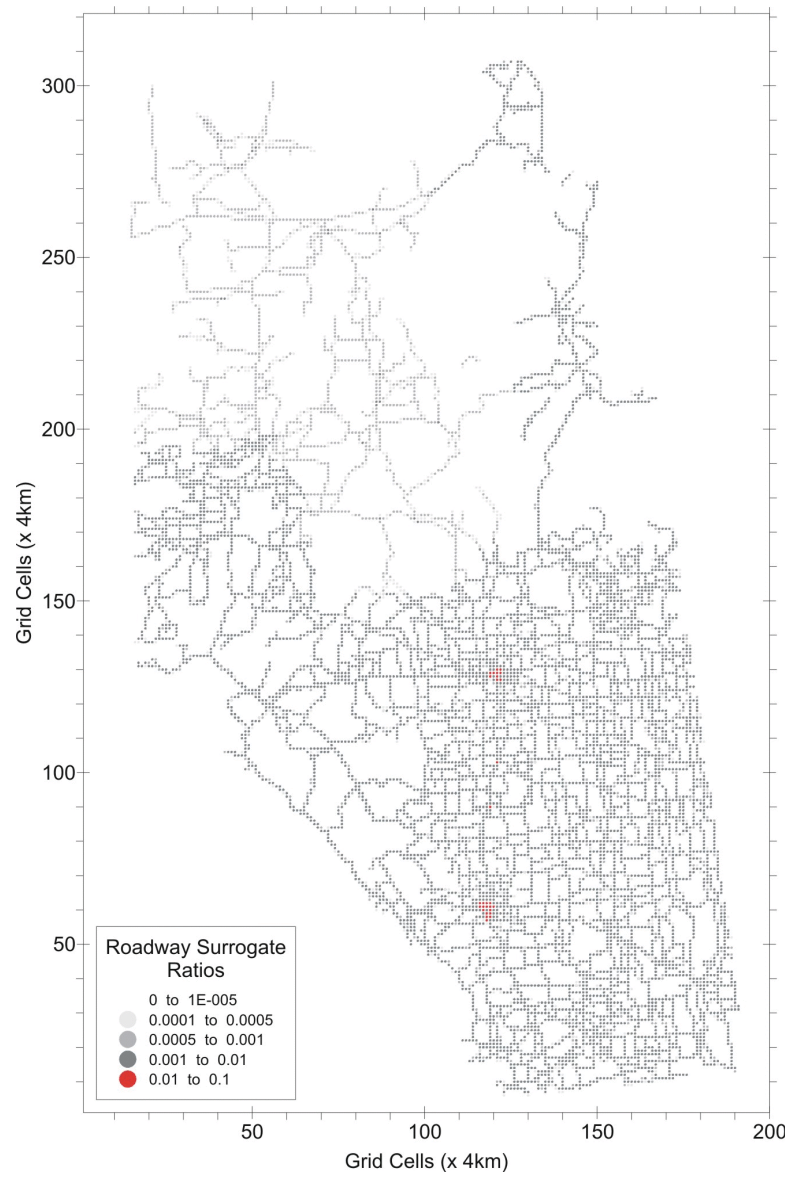


Figure 6. Gridded Alberta annual NO_x emissions. **Figure 7.** Gridded Alberta annual SO_x emissions.

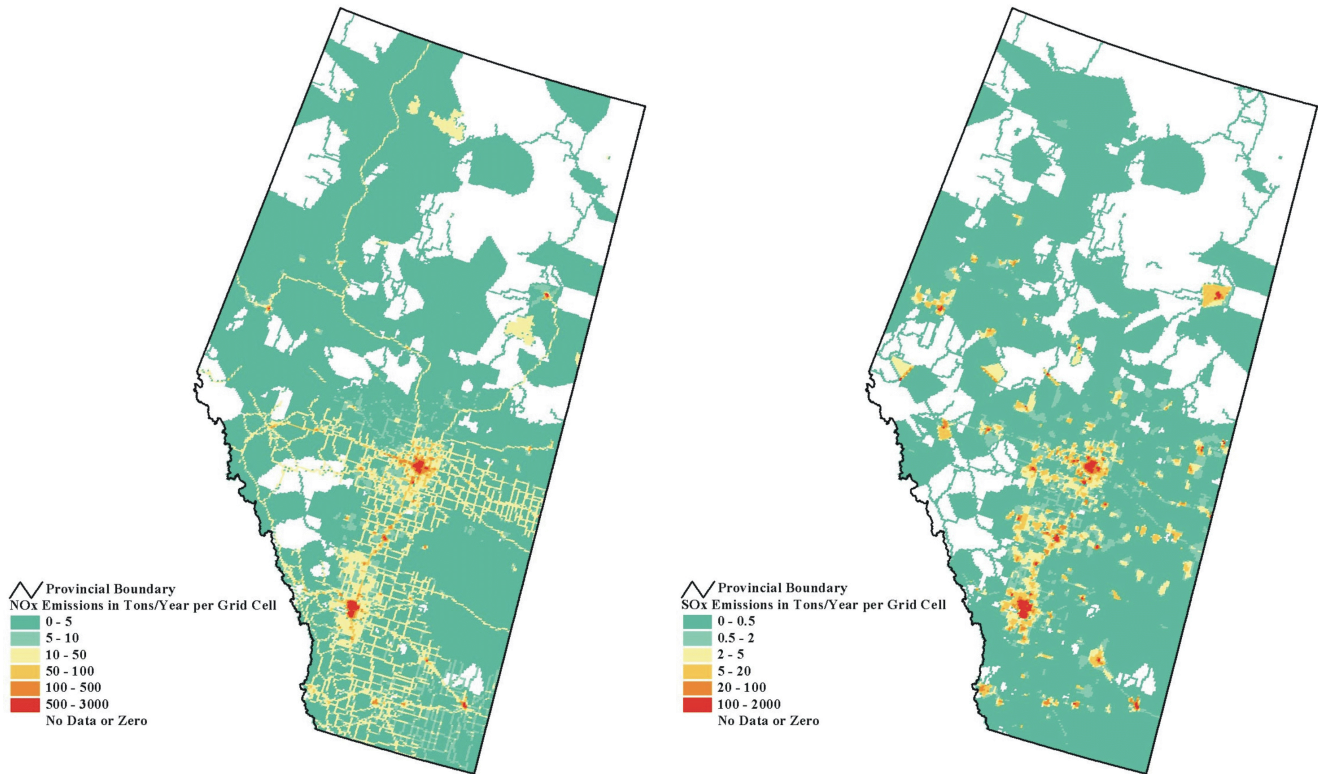


Figure 8. GTAA study area and digital road network.

